Exploiter la compétence de couples de cas pour améliorer le raisonnement à partir de cas par analogie

Jean Lieber[1] **Emmanuel Nauer**[1] **Henri Prade**[2]

[1] Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France
[2] IRIT, CNRS & Université de Toulouse, 31062 Toulouse cedex 9, France
Jean.Lieber@loria.fr Emmanuel.Nauer@loria.fr Henri.Prade@irit.fr

Résumé

Une proportion analogique est une relation quaternaire qui se lit « *a* est à *b* ce que *c* est à *d* » et qui vérifie certaines propriétés de symétrie et de permutation. Cette relation, qui met en œuvre deux couples, est le fondement de l'approche de raisonnement à partir de cas appelée extrapolation analogique, qui consiste à remémorer trois cas formant une proportion analogique avec le problème à résoudre dans l'espace des problèmes, puis à trouver une solution à ce problème en résolvant une équation analogique dans l'espace des solutions. Ce travail étudie comment la notion de compétence des couples de cas sources peut être évaluée et utilisée pour améliorer l'extrapolation. Un prétraitement de la base de cas associe à chaque couple de cas une compétence sous la forme de deux scores : le support et la confiance du couple de cas, calculé à partir d'autres couples avec qui il forme une proportion analogique. Une évaluation dans un cadre booléen montre que l'utilisation des compétences des couples de cas améliore significativement les résultats d'un système de raisonnement à partir de cas par analogie.

Mots clefs : compétence, extrapolation, inférence analogique, proportion analogique, raisonnement à partir de cas

Abstract

An analogical proportion is a quaternary relation that is to be read "*a* is to *b* as *c* is to *d*", verifying some symmetry and permutation properties. As can be seen, it involves a pair of pairs. Such a relation is at the basis of an approach to case-based reasoning called analogical extrapolation, which consists in retrieving three cases forming a proportional analogy with the target problem in the problem space and then in finding a solution to this problem by solving an analogical equation in the solution space. This paper studies how the notion of competence of pairs of source cases can be estimated and used in order to improve extrapolation. A preprocessing of the case base associates to each case pair a competence given by two scores: the support and the confidence of the case pair, computed on the basis of other case pairs forming a proportional analogy with it. An evaluation in a Boolean setting shows that using case pair competences improves significantly the result of the analogical extrapolation process.

Keywords: analogical inference, analogical proportion, case-based reasoning, competence, extrapolation

1 Introduction

Dans un travail récent [13], les auteurs ont montré que le raisonnement à partir de cas (RàPC [16]) peut ne pas être fondé uniquement sur un raisonnement utilisant la similarité (à la recherche des cas résolus les plus proches) mais peut également utiliser des proportions analogiques à des fins d'extrapolation. L'extrapolation repose sur l'inférence analogique qui utilise des triplets de cas pour construire la solution d'un quatrième (nouveau) cas par le biais d'un mécanisme d'adaptation.

Habituellement, plusieurs triplets de la base de cas peuvent être utilisés pour prédire la solution du quatrième cas et les prédictions peuvent diverger. En réalité, il a été démontré que dans un contexte booléen une inférence de ce type ne produit pas d'erreurs (i.e. tous les triplets conduisent vers la même prédiction) si et seulement si la fonction qui associe une solution à la description du cas est une fonction booléenne affine [4]. C'est pourquoi, lorsque la fonction n'est pas supposée affine, une procédure de vote est mise en œuvre sur les prédictions des triplets.

Une telle procédure est assez basique et n'exploite pas pleinement la base de cas. En effet, il se peut que certains triplets dans la base échouent dans la prédiction de la bonne réponse d'un autre cas déjà présent dans la base, traité comme un nouveau cas à résoudre. Dans ce travail, nous proposons de prendre en compte ce type d'information pour restreindre le nombre de triplets utilisés pour effectuer une prédiction.

Ce document est organisé de la façon suivante. La section 2 fournit les prérequis nécessaires aux proportions analogiques et aux notations sur le RàPC utilisées dans ce document. La section 3 explique comment restreindre les triplets autorisés à participer à une prédiction donnée. La section 4 présente une expérimentation qui montre le gain en précision de la nouvelle procédure d'inférence. Enfin, la section 5 situe ce travail au regard de l'existant.

2 Préliminaires

Cette section présente dans un premier temps le cadre formel de ce travail : celui des représentations nominales et, plus particulièrement, les représentations par n-uplets de booléens. Puis, elle rappelle quelques notions et notations, d'abord sur les proportions analogiques puis sur le raisonnement à partir de cas.

2.1 Représentation nominale et contexte booléen

Les représentations de type attribut-valeur sont souvent utilisées en RàPC (voir, e.g., [11]). Une représentation nominale est une représentation de type attribut-valeur où le co-domaine de l'attribut est fini (et, typiquement, petit). Plus formellement, soit $\mathcal{U}_1$, $\mathcal{U}_2$, ..., $\mathcal{U}_p$ p ensembles finis et $\mathcal{U} = \mathcal{U}_1 \times \mathcal{U}_2 \times \ldots \times \mathcal{U}_p$. Un attribut sur $\mathcal{U}$ est une des p projections $(x_1, x_2, \ldots, x_p) \in \mathcal{U} \mapsto x_i \in \mathcal{U}_i$ ($i \in \{1, 2, \ldots, p\}$).

Une représentation booléenne est une représentation nominale où $\mathcal{U}_1 = \mathcal{U}_2 = \ldots = \mathcal{U}_p = \mathbb{B}$, avec $\mathbb{B} = \{0, 1\}$ l'ensemble des valeurs booléennes : la valeur « faux » est assimilée à l'entier 0, et la valeur « vrai » est assimilée à 1. Les opérateurs booléens $\neg$, $\land$ et $\lor$ sont définis, pour $\mathsf{a}, \mathsf{b} \in \mathbb{B}$, par $\neg\mathsf{a} = 1 - \mathsf{a}$, $\mathsf{a} \land \mathsf{b} = \begin{cases} 1 & \text{if } \mathsf{a} = \mathsf{b} = 1 \\ 0 & \text{sinon} \end{cases}$, $\mathsf{a} \lor \mathsf{b} = \neg(\neg\mathsf{a} \land \neg\mathsf{b})$, $\mathsf{a} \equiv \mathsf{b} = (\neg\mathsf{a} \lor \mathsf{b}) \land (\neg\mathsf{b} \lor \mathsf{a})$. Un élément de $\mathbb{B}^p$ est noté sans virgules ni parenthèses, par exemple 01101 correspond à $(0, 1, 1, 0, 1)$.

2.2 Proportions analogiques

Étant donné un ensemble $\mathcal{U}$, une proportion analogique sur $\mathcal{U}$ est une relation quaternaire sur $\mathcal{U}$, notée $a\!:\!b\!::\!c\!:\!d$ pour $(a, b, c, d) \in \mathcal{U}^4$, et satisfiant les postulats suivants (pour $a, b, c, d \in \mathcal{U}$) :

(Réflexivité) $a\!:\!b\!::\!a\!:\!b$.

(Symétrie) Si $a\!:\!b\!::\!c\!:\!d$ alors $c\!:\!d\!::\!a\!:\!b$.

(Échange des moyens) Si $a\!:\!b\!::\!c\!:\!d$ alors $a\!:\!c\!::\!b\!:\!d$.

Étant donné un ensemble fini $\mathcal{U}_i$, la relation définie ci-dessous est une proportion analogique [15] :

$$a\!:\!b\!::\!c\!:\!d \stackrel{\text{déf}}{=} ((a \equiv b) \land (c \equiv d)) \lor ((a \equiv c) \land (b \equiv d))$$

Par conséquent, dans le cas nominal, les quadruplets (a, b, c, d) en analogie sont d'une des trois formes suivantes : (s, s, s, s), (s, t, s, t) et (s, s, t, t) pour $s, t \in \mathcal{U}_i$. En particulier, si $\mathcal{U}_i = \mathbb{B}$, l'ensemble des $(\mathsf{a}, \mathsf{b}, \mathsf{c}, \mathsf{d}) \in \mathbb{B}^4$ tels que $\mathsf{a\!:\!b\!::\!c\!:\!d}$ est donc $\{0000, 1111, 0011, 1100, 0101, 1010\}$.

Pour un $\mathcal{U} = \mathcal{U}_1 \times \mathcal{U}_2 \times \ldots \times \mathcal{U}_p$ fini, la proportion analogique suivante peut être définie :

$$\begin{aligned} a\!:\!b\!::\!c\!:\!d = {} & a_1\!:\!b_1\!::\!c_1\!:\!d_1 \\ & \land a_2\!:\!b_2\!::\!c_2\!:\!d_2 \\ & \land \ldots \\ & \land a_p\!:\!b_p\!::\!c_p\!:\!d_p \end{aligned}$$

Étant donné $a, b, c \in \mathcal{U}$, résoudre l'*équation analogique* $a\!:\!b\!::\!c\!:\!y$ consiste à trouver les $y \in \mathcal{U}$ tels que cette relation soit vérifiée. Dans une représentation nominale, une telle équation a 0 ou 1 solution. Plus précisément :

- Si $a = b$, la solution est $y = c$.
- Si $a = c$, la solution est $y = b$.
- Sinon, $a\!:\!b\!::\!c\!:\!y$ n'a pas de solution.

2.3 Notations et hypothèses du RàPC

Soient $\mathcal{P}$ et $\mathcal{S}$ deux ensembles. Un *problème* $\mathtt{x}$ est par définition un élément de $\mathcal{P}$ et une *solution* $\mathtt{y}$, un élément de $\mathcal{S}$. Si $a \in \mathcal{P} \times \mathcal{S}$, alors $\mathtt{x}^a$ et $\mathtt{y}^a$ sont respectivement sa partie problème et sa partie solution : $a = (\mathtt{x}^a, \mathtt{y}^a)$. Soit $\rightsquigarrow$ une relation sur $\mathcal{P} \times \mathcal{S}$. Pour $(\mathtt{x}, \mathtt{y}) \in \mathcal{P} \times \mathcal{S}$, $\mathtt{x} \rightsquigarrow \mathtt{y}$ exprime « $\mathtt{x}$ a pour solution $\mathtt{y}$ » ou « $\mathtt{y}$ résout $\mathtt{x}$ ». Un *cas* est un couple $(\mathtt{x}, \mathtt{y})$ tel que $\mathtt{x} \rightsquigarrow \mathtt{y}$. Le but d'un système de RàPC est de résoudre des problèmes, i.e., il doit approcher la relation $\rightsquigarrow$: étant donné $\mathtt{x}^{\mathtt{cible}} \in \mathcal{P}$ (le problème *cible*), il doit proposer $\mathtt{y}^{\mathtt{cible}} \in \mathcal{S}$ tel qu'il est plausible que $\mathtt{x}^{\mathtt{cible}} \rightsquigarrow \mathtt{y}^{\mathtt{cible}}$. Pour cela, un ensemble fini de cas, appelé la base de cas, notée BC, est utilisée. Un élément de BC est appelé un *cas source*. Hormis la base de cas, d'autres *conteneurs* de connaissances sont souvent utilisés [16], mais ne sont pas considérés dans ce travail.

La façon habituelle de définir le processus de RàPC consiste à sélectionner un ensemble de k cas sources en relation avec $\mathtt{x}^{\mathtt{cible}}$ (étape de remémoration) et puis à résoudre $\mathtt{x}^{\mathtt{cible}}$ à l'aide des cas remémorés (étape d'adaptation). D'autres étapes sont considérées dans le modèle classique de RàPC [1], mais pas dans ce travail. Dans [13], trois approches sont présentées pour $k \in \{1, 2, 3\}$, correspondant à trois relations entre k-uplets de cas sources et le problème à résoudre. Dans ce travail, l'approche avec $k = 3$ est rappelée dans la section suivante.

3 Améliorer l'extrapolation grâce à la compétence des couples de cas

Cette section présente l'approche proposée. Dans un premier temps, on montrera comment une notion de compétence associée aux couples de cas peut être utilisée pour améliorer l'extrapolation. Puis, la notion de compétence est définie formellement. Enfin, des stratégies utilisant cette notion en vue d'améliorer l'extrapolation sont décrites.

3.1 Principes

Le principe de l'inférence par proportion analogique [18] peut être énoncé comme suit (avec l'utilisation des notations de RàPC introduites précédemment ; $a = (\mathtt{x}^a, \mathtt{y}^a)$, $b = (\mathtt{x}^b, \mathtt{y}^b)$, $c = (\mathtt{x}^c, \mathtt{y}^c)$ et $d = (\mathtt{x}^d, \mathtt{y}^d)$ sont quatre cas) :

$$\frac{\mathtt{x}^a : \mathtt{x}^b :: \mathtt{x}^c : \mathtt{x}^d \text{ est vérifiée}}{\mathtt{y}^a : \mathtt{y}^b :: \mathtt{y}^c : \mathtt{y}^d \text{ est vérifiée}}$$

Résoudre un nouveau problème $\mathtt{x}^{\mathtt{cible}}$, consiste à considérer tous les triplets de cas sources (a, b, c) tels que $\mathtt{x}^a : \mathtt{x}^b :: \mathtt{x}^c : \mathtt{x}^{\mathtt{cible}}$ est vérifiée et tels que l'équation $\mathtt{y}^a : \mathtt{y}^b :: \mathtt{y}^c : y$ est résoluble. Soit $\mathcal{T}$ l'ensemble de tous ces triplets. L'implantation de cette inférence utilise un vote parmi tous les triplets de $\mathcal{T}$ et choisit la solution y trouvée sur le plus grand ensemble de triplets. Ceci est le principe de l'extrapolation analogique (ou, simplement, extrapolation [13]).

Dans la suite, nous supposons pour simplifier que tous les attributs sont nominaux (p. ex., booléens). Lorsqu'il y a un seul attribut pour les solutions, la tâche de résolution de problème est une tâche de classification (consistant à trouver la classe $\mathtt{y}^{\mathtt{cible}} \in \mathcal{S}$ devant être associée à $\mathtt{x}^{\mathtt{cible}}$). Lorsqu'il y a plusieurs attributs, il est possible de les traiter un par un si et seulement si ils sont logiquement indépendants. Sinon, le vote doit être organisé entre les vecteurs complets décrivant les différentes solutions. Dans la suite, nous supposons l'indépendance et nous considérons une des composantes $\mathtt{y}_i$ d'une solution $\mathtt{y}$ (ainsi, l'index i est inutile : $\mathtt{y}_i$ est noté par $\mathtt{y}$).

Cependant, on peut se demander si tous les triplets de $\mathcal{T}$ impliqués dans un vote pour faire une prédiction particulière ont la même légitimité. En effet, des leçons peuvent être tirées de $\mathcal{T}$ en observant lors d'une prédiction de solution pour un problème déjà présent dans $\mathcal{T}$ s'il existe des triplets produisant une prédiction erronée, comme suggéré dans [14]. La situation peut être mieux analysée en termes de couples, comme on le montre à présent.

La table 1 donne un exemple avec trois couples de booléens tels que $\mathtt{a}:\mathtt{b}::\mathtt{c}:\mathtt{d}$ et $\mathtt{a}':\mathtt{b}'::\mathtt{c}:\mathtt{d}$ sont vérifiées dans toutes les colonnes, exceptée la dernière colonne, 'S', colonne solution. Les deux premières colonnes 'D' (comme désaccord) montrent les motifs possibles exprimant que $\mathtt{a}$ et $\mathtt{b}$ diffèrent de la même façon que $\mathtt{c}$ et $\mathtt{d}$ et que $\mathtt{a}'$ et $\mathtt{b}'$([1]). Les colonnes 'A' (comme accord) montrent toutes les possibilités où $\mathtt{a}$ et $\mathtt{b}$ sont en accord, avec $\mathtt{c}$ et $\mathtt{d}$ également en accord, ainsi que $\mathtt{a}'$ et $\mathtt{b}'$, mais éventuellement d'une façon différente [2]. Si on regarde la valeur de $\mathtt{d}$ dans la colonne 'S' et que l'on cherche à la prédire à partir des autres valeurs de cette colonne, l'équation $\mathtt{a}:\mathtt{b}::\mathtt{c}:y$ donne le bon résultat, tandis que l'équation $\mathtt{a}':\mathtt{b}'::\mathtt{c}:y$ donne le mauvais.

Ainsi, pour chaque couple comme $(\mathtt{a}, \mathtt{b})$ ou $(\mathtt{a}', \mathtt{b}')$ dans la table, il est possible de compter le nombre de fois où les couples conduisent à une bonne ou à une mauvaise prédiction pour un exemple pris dans BC (via des $\mathtt{c}$ appropriés également présents dans BC). Cela fournit un moyen de favoriser des triplets contenant des couples menant à de bonnes prédictions, dans la procédure de vote.

L'idée ci-dessus, de regarder les couples de cas, peut être mise en relation avec l'interprétation d'un couple de cas (a, b) comme une règle potentielle exprimant soit que le changement de $\mathtt{x}^a$ à $\mathtt{x}^b$ explique le changement de $\mathtt{y}^a$ à $\mathtt{y}^b$, indépendamment du contexte (encodé par les attributs pour lesquelles les exemples sont en accord), soit que le changement de $\mathtt{x}^a$ à $\mathtt{x}^b$ ne modifie pas la solution (si $\mathtt{y}^a = \mathtt{y}^b$). Cette vision des couples de cas comme des règles a déjà été proposée en RàPC pour trouver des règles d'adaptation [9, 5, 6] et par la suite dans une perspective d'inférence fondée sur les proportions analogiques [3, 2].

Par conséquent, nous sommes intéressés par un prétraitement, afin de découvrir des *ruptures d'analogie* dans $\mathcal{T}$. Par rupture d'analogie, nous entendons l'existence d'un quadruplet de cas (a, b, c, d) tel que (i) $\mathtt{x}^a : \mathtt{x}^b :: \mathtt{x}^c : \mathtt{x}^d$ est vérifié, tandis que (ii) $\mathtt{y}^a : \mathtt{y}^b :: \mathtt{y}^c : \mathtt{y}^d$ *n'est pas* vérifié. Si des ruptures d'analogie peuvent être trouvées dans $\mathcal{T}$, cela signifie que la fonction booléenne partiellement inconnue qui associe à un problème sa solution (ou sa classe) ne peut pas être affine [4]. Dans cette situation, une inférence analogique ne peut pas être appliquée aveuglément avec n'importe quel triplet, et nous devrions tenir compte de ces ruptures d'analogie, en introduisant des restrictions supplémentaires pour le choix de triplets appropriés.

Plus précisément, l'idée est de faire un prétraitement des couples $(a, b) \in \mathtt{BC}^2$, en associant à chacun d'eux une *compétence*. L'intuition derrière cette notion est que plus un couple de cas est compétent pour résoudre des problèmes, plus il peut jouer un rôle dans le processus de vote. Pour évaluer la compétence d'un couple $(a, b) \in \mathtt{BC}^2$, il doit être mis en proportion analogique avec d'autres couples $(c, d) \in \mathtt{BC}^2$ tels que le triplet (a, b, c) peut être utilisé pour résoudre le problème $\mathtt{x}^d$ par extrapolation. Si le résultat de l'extrapolation y est égal à $\mathtt{y}^d$, la compétence du couple de cas (a, b) augmente, sinon elle diminue. La définition de la

1. $D(0/1)$ indique le désaccord entre $\mathtt{a}$ et $\mathtt{b}$ (resp., entre $\mathtt{c}$ et $\mathtt{d}$ et entre $\mathtt{a}'$ et $\mathtt{b}'$) où le premier vaut 0 et le second vaut 1. $D(1/0)$ est le désaccord inverse.

2. Dans la table 1, $A(u, v, w)$ signifie que $\mathtt{a} = \mathtt{b} = u$, $\mathtt{c} = \mathtt{d} = v$ et $\mathtt{a}' = \mathtt{b}' = w$.

	$D(0/1)$	$D(1/0)$	$A(0,0,0)$	$A(0,0,1)$	$A(0,1,0)$	$A(0,1,1)$	$A(1,0,0)$	$A(1,0,1)$	$A(1,1,0)$	$A(1,1,1)$	S
a	0	1	0	0	0	0	1	1	1	1	0
b	1	0	0	0	0	0	1	1	1	1	1
c	0	1	0	0	1	1	0	0	1	1	0
d	1	0	0	0	1	1	0	0	1	1	1
a′	0	1	0	1	0	1	0	1	0	1	0
b′	1	0	0	1	0	1	0	1	0	1	0

TABLE 1 – Double appariement de couples (a, b), (c, d) et $(\mathsf{a}', \mathsf{b}')$: rupture d'analogie en S.

compétence est détaillée dans la section suivante.

La compétence d'un couple de cas peut être utilisée au moment de la résolution d'un nouveau problème, avec différentes stratégies. La section 3.3 présente certaines de ces stratégies qui sont évaluées expérimentalement dans un cadre booléen en section 4.

3.2 Acquisition des compétences des couples de cas

Soit (a, b) un couple de cas sources : $a = (\mathtt{x}^a, \mathtt{y}^a) \in \mathtt{BC}$ et $b = (\mathtt{x}^b, \mathtt{y}^b) \in \mathtt{BC}$. La *compétence* du couple (a, b) est définie par deux scores : le support et la confiance de (a, b), définis ci-dessous à partir du principe présenté précédemment.

Tout d'abord, soit $\mathtt{RésolublesPar}(a, b)$ l'ensemble des couples de cas sources $(c, d) \neq (a, b)$ tels que le triplet (a, b, c) peut être utilisé pour résoudre $\mathtt{x}^d$ par extrapolation : $\mathtt{x}^a : \mathtt{x}^b :: \mathtt{x}^c : \mathtt{x}^d$ avec l'équation $\mathtt{y}^a : \mathtt{y}^b :: \mathtt{y}^c : y$ résoluble (et ainsi, sa solution y est unique dans une représentation nominale). Formellement :

$$\mathtt{RésolublesPar}(a, b) = \left\{ (c, d) \in \mathtt{BC}^2 \;\middle|\; \begin{array}{l} (c, d) \neq (a, b), \\ \mathtt{x}^a : \mathtt{x}^b :: \mathtt{x}^c : \mathtt{x}^d \\ \text{et l'équation} \\ \mathtt{y}^a : \mathtt{y}^b :: \mathtt{y}^c : y \\ \text{a une solution} \end{array} \right\}$$

En d'autres termes, $\mathtt{RésolublesPar}(a, b)$ est l'ensemble des couples de cas sources (c, d) tels que c peut être adapté en une solution de $\mathtt{x}^d$ en utilisant (a, b) comme une règle d'adaptation (sans considérer le cas trivial $(a, b) = (c, d)$). Le support de (a, b), $\mathtt{supp}(a, b)$, est simplement le nombre de tels couples :

$$\mathtt{supp}(a, b) = |\mathtt{RésolublesPar}(a, b)|$$

Parmi les $(c, d) \in \mathtt{RésolublesPar}(a, b)$ certains conduisent à la bonne solution ($y = \mathtt{y}^d$) et d'autres non. Les premiers constituent l'ensemble suivant :

$$\begin{aligned} &\mathtt{RésolublesCorrectementPar}(a, b) \\ &= \left\{ (c, d) \in \mathtt{RésolublesPar}(a, b) \mid \mathtt{y}^a : \mathtt{y}^b :: \mathtt{y}^c : \mathtt{y}^d \right\} \end{aligned}$$

Par exemple, si $\mathtt{supp}(a, b) = 6$ et $|\mathtt{RésolublesCorrectementPar}(a, b)| = 4$, cela signifie que (a, b), considéré comme une règle, a été testé sur la base de cas 6 fois et a donné 4 bonnes réponses. La proportion de bonnes réponses est donc $4/6 = 2/3$. Cette proportion est appelée confiance de (a, b). Un cas particulier doit être considéré quand $\mathtt{supp}(a, b) = 0$. Cela signifie que « la règle d'adaptation » (a, b) ne peut pas être testée sur la base de cas. Dans ce cas, la valeur de la confiance est fixé à 0,5 (meilleure qu'une confiance de, par exemple, $3/7$ pour laquelle la règle échoue plus souvent qu'elle ne réussit et moins bien qu'une confiance de $4/7$ pour laquelle la règle donne de bonnes réponses plus souvent que de mauvaises). En résumé, la confiance d'un couple (a, b) est :

$$\begin{aligned} &\mathtt{conf}(a, b) \\ &= \begin{cases} \frac{|\mathtt{RésolublesCorrectementPar}(a,b)|}{\mathtt{supp}(a,b)} & \text{si } \mathtt{supp}(a, b) \neq 0 \\ 0{,}5 & \text{sinon} \end{cases} \end{aligned}$$

3.3 Utilisation des compétences de couples de cas dans des stratégies de sélection et de vote

Étant donné un problème cible $\mathtt{x}^{\mathtt{cible}}$, l'extrapolation consiste à remémorer des triplets $(a, b, c) \in \mathtt{BC}^3$ tels que $\mathtt{x}^a : \mathtt{x}^b :: \mathtt{x}^c : \mathtt{x}^{\mathtt{cible}}$ et à adapter ces triplets en résolvant l'équation $\mathtt{y}^a : \mathtt{y}^b :: \mathtt{y}^c : y$ pour chacun de ces triplets (évidemment, les triplets (a, b, c) pour lesquelles l'équation n'a pas de solution ne sont pas considérés). Ainsi, le résultat de l'extrapolation est l'ensemble $\mathcal{R}$ des $((a, b, c), y) \in \mathtt{BC}^3 \times \mathcal{S}$, y étant le résultat de l'extrapolation de (a, b, c) afin de résoudre $\mathtt{x}^{\mathtt{cible}}$. Maintenant, la question est comment considérer toutes ces solutions y pour proposer une unique solution $\mathtt{y}^{\mathtt{cible}}$ pour $\mathtt{x}^{\mathtt{cible}}$. Pour cela, quatre stratégies sont considérées.

La première, appelée `sansComp`, effectue simplement un vote avec toutes les valeurs de y, indépendamment des compétences. La solution proposée est ainsi

$$\mathtt{y}^{\mathtt{cible}} = \underset{\widehat{y}}{\operatorname{argmax}} \left| \{ ((a, b, c), y) \in \mathcal{R} \mid y = \widehat{y} \} \right|$$

C'est la stratégie utilisée dans [13] et elle constitue une référence pour l'évaluation.

La deuxième stratégie, appelée `toutesConf`, considère tous les $((a, b, c), y) \in \mathcal{R}$ et effectue un vote pondéré par la confiance :

$$\mathtt{y}^{\mathtt{cible}} = \underset{\widehat{y}}{\operatorname{argmax}} \sum_{((a,b,c),y) \in \mathcal{R}, y = \widehat{y}} \mathtt{conf}(a, b)$$

La troisième stratégie, appelée $\texttt{maxConf}$, considère uniquement les $((a,b,c),y) \in \mathcal{R}$ ayant la plus haute confiance, et effectue un vote à partir de ces dernières. Formellement :

$$\begin{aligned}\text{avec } \texttt{conf}_{\max} &= \max\{\texttt{conf}(a,b) \mid ((a,b,c),y) \in \mathcal{R}\} \\ \text{et } \mathcal{R}^* &= \{((a,b,c),y) \in \mathcal{R} \mid \texttt{conf}(a,b) = \texttt{conf}_{\max}\} \\ \mathtt{y}^{\texttt{cible}} &= \operatorname*{argmax}_{\widehat{y}} \left|\{((a,b,c),y) \in \mathcal{R}^* \mid y = \widehat{y}\}\right| \qquad (1)\end{aligned}$$

La quatrième stratégie, appelée $\texttt{maxConfSupp}$, est similaire à la troisième, excepté qu'elle utilise à la fois la confiance et le support pour établir une préférence. Plus précisément, cette stratégie s'appuie sur un ordre de préférence des couples de cas défini ci-dessous (pour $(a,b),(a',b') \in \texttt{BC}^2$:

$$(a,b) \succcurlyeq (a',b') \quad \text{si} \quad \left| \begin{array}{l} \texttt{conf}(a,b) > \texttt{conf}(a',b') \ \textbf{ou} \\ (\texttt{conf}(a,b) = \texttt{conf}(a',b') \\ \quad \textbf{et}\ \texttt{supp}(a,b) \geqslant \texttt{supp}(a',b')) \end{array} \right.$$

En d'autres termes, la confiance est le critère principal mais, en cas d'égalité, plus le support est élevé, plus le couple de cas (a,b) est considéré comme étant compétent. Par exemple, si $\texttt{conf}(a,b) = \texttt{conf}(a',b') = 0{,}75$, $\texttt{supp}(a,b) = 8$ et $\texttt{supp}(a',b') = 4$, alors (a,b) donne la bonne réponse dans 6 situations sur 8, tandis que (a',b') le fait dans 3 situations sur 4. Soit alors $\mathcal{R}^*$ l'ensemble des $((a,b,c),y) \in \mathcal{R}$ tels que (a,b) est maximal pour $\succcurlyeq$. Alors, $\mathtt{y}^{\texttt{cible}}$ résulte d'un vote, comme décrit précédemment dans l'équation (1).

L'intérêt de considérer un triplet (a,b,c) dans la procédure de vote à la fin du processus d'inférence est évalué en terme de compétence du couple (a,b). Comme les proportions analogiques sont stables pour la permutation centrale, on peut également penser à considérer les couples (a,c). Des investigations préliminaires, utilisant différentes combinaisons (minimum, maximum, somme et produit) des confiances de (a,b) et (a,c) n'ont pas montré d'apport significatif par rapport à l'utilisation seule de la confiance de (a,b) ; c'est pourquoi nous nous sommes restreints, dans ce travail, à ce dernier type de prise en compte de la compétence. Cependant, la combinaison n'a porté que sur l'utilisation de la confiance brute de tous les triplets (comme dans la méthode $\texttt{toutesConf}$) sans considérer de stratégie de restriction des triplets participant au vote (comme dans les stratégies $\texttt{maxConf}$ et $\texttt{maxConfSupp}$). L'étude plus approfondie d'une stratégie de combinaison reposant sur la confiance mais également sur le support constitue une perspective de ce travail.

4 Évaluation

L'objectif de l'évaluation est d'étudier l'impact des stratégies pour la sélection des couples de cas participant au vote sur différents types de fonctions booléennes.

4.1 Contexte expérimental

Dans nos expériences, $\mathcal{P} = \mathbb{B}^8$ et $\mathcal{S} = \mathbb{B}$. Ici $\rightsquigarrow$ est supposé être fonctionnelle : $\rightsquigarrow = \mathtt{f}$.

La fonction $\mathtt{f}$ est générée de manière aléatoire à l'aide des générateurs suivants qui sont fondées sur deux des formes normales les plus classiques dans le but d'avoir une diversité de fonctions générées :

DNF $\mathtt{f}$ est générée en forme normale disjonctive, c.-à-d., que $\mathtt{f}(\mathtt{x})$ est une disjonction de n_{disj} clauses (conjonctions de littéraux), par exemple :
$\mathtt{f}(\mathtt{x}) = (\mathtt{x}_1 \wedge \neg\mathtt{x}_7) \vee (\neg\mathtt{x}_3 \wedge \mathtt{x}_7 \wedge \mathtt{x}_8) \vee \mathtt{x}_4$.
La valeur de n_{disj} est choisie au hasard uniformément dans $\{3,4,5\}$. Chaque disjonction est générée sur la base de deux paramètres, $p^+ > 0$ et $p^- > 0$, avec $p^+ + p^- < 1$: dans chaque clause, pour chaque variable x_i, elle apparaît sous forme de littéral positif (resp., négatif) avec une probabilité p^+ (resp., p^-). Dans nos expériences, nous avons choisi $p^+ = p^- = 0{,}1$([3]).

Pol $\mathtt{f}$ est générée sous forme polynomiale : $\mathtt{f}$ a la même forme que si elle était générée par DNF, sauf que les disjonctions ($\vee$) sont remplacées par des ou exclusifs ($\oplus$). Comme seuls les littéraux positifs apparaissent dans la forme normale polynomiale, le paramètre $p^- = 0$.

La base de cas BC est générée aléatoirement, avec des valeurs de taille $|\texttt{BC}| \in \{32, 64, 96, 128\}$, i.e. $|\texttt{BC}|$ est compris entre $\frac{1}{8}$ et $\frac{1}{2}$ de $|\mathcal{P}| = 2^8 = 256$. Chaque cas source $(\mathtt{x},\mathtt{y})$ est généré comme suit : $\mathtt{x}$ est choisi au hasard dans $\mathcal{P}$ avec une distribution uniforme et $\mathtt{y} = \mathtt{f}(\mathtt{x})$.

Soient `#pb_cibles` le nombre de problèmes cibles soumis au système, `#rep` le nombre de réponses (correctes ou incorrectes) (`#pb_cibles − #rep` est le nombre de problèmes cibles pour lesquels le système échoue à proposer une solution), et `#rep_ok` le nombre de réponses correctes. Pour chaque stratégie de sélection et vote, les scores suivants sont calculés :

Le taux d'erreur `%err` est la moyenne des $\left(1 - \dfrac{\texttt{\#rep\_ok}}{\texttt{\#rep}}\right) \times 100 \in [0, 100]$.

Le taux de réponse `%rep` est la moyenne des rapports $\dfrac{\texttt{\#rep}}{\texttt{\#pb\_cibles}} \times 100 \in [0, 100]$. Si le système retourne toujours une réponse (correcte ou non) alors `%rep` = 100.

3. Un générateur CNF, générant des formules en CNF (forme normale conjonctive : conjonction de disjonctions de littéraux) peut également être considéré. Cependant, ceci n'apporterait rien car ce serait dual avec le générateur DNF pour deux raisons. Premièrement, les inférences effectuées sont indépendantes du codage, signifiant que remplacer des attributs par leur négations ne changent pas le résultat de l'inférence, en particulier, pour $\mathtt{a},\mathtt{b},\mathtt{c},\mathtt{d} \in \mathbb{B}$, $\mathtt{a}:\mathtt{b}::\mathtt{c}:\mathtt{d}$ si et seulement si $\neg\mathtt{a}:\neg\mathtt{b}::\neg\mathtt{c}:\neg\mathtt{d}$. Deuxièmement, si $\mathtt{f}$ est obtenu à partir d'un générateur DNF alors $\neg\mathtt{f}$ peut facilement être transformé en une fonction g écrite en CNF en utilisant les lois de De Morgan, et la distribution de g obtenue de cette façon serait la même que la distribution d'un générateur CNF avec les mêmes paramètres.

La moyenne est calculée sur 1 million de résolutions de problème pour chaque générateur, nécessitant la génération de 1420 `f` pour chacun d'eux. La moyenne du temps de calcul d'une session de RàPC (remémoration et adaptation pour résoudre un problème) est d'environ 2 ms sur un ordinateur portable courant.

Par souci de reproductibilité, le code des expérimentations est disponible sur `https://tinyurl.com/analogyCBRTests`, avec les résultats détaillés (fonctions générées et détails des évaluations).

4.2 Résultats

La table 2 présente les taux d'erreur et de réponse pour les différentes stratégies de sélection et de vote pour les deux générateurs, avec une application à différentes tailles de la base de cas. Les courbes du taux d'erreur sont données en Figure 1.

Étant donnée un générateur de fonction et une taille de base de cas, le taux de réponse est le même pour chacune des quatre stratégies car chacune des stratégies de sélection fournit des résultats pour un problème qui peut être résolu par l'approche sans sélection `sansComp` (i.e. si un triplet a été trouvé pour résoudre un cas par `sansComp`, ce triplet sera considéré par les trois stratégies de sélection et soit il participera à la résolution de $\mathtt{x}^{\mathtt{cible}}$, soit il existe un autre triplet qui est « meilleur » vis-à-vis de la procédure de sélection). Le taux de réponse est élevé pour toutes les méthodes : plus de 96% pour $|\mathtt{BC}| = 32$ et 100% pour $|\mathtt{BC}| \geqslant 64$.

Excepté pour $|\mathtt{BC}| = 32$, qui semble être un ensemble d'apprentissage trop petit pour calculer les compétences, le taux d'erreur montre que l'hypothèse de sélection des couples de cas améliore la précision. Pour les deux générateurs, toutes les stratégies de sélection donnent de meilleurs résultats que la stratégie de référence (`sansComp`). Cependant, l'amélioration est assez différente en fonction de la stratégie de sélection : plus la sélection des couples de cas est contrainte, plus le taux d'erreur décroit. `toutesConf` diminue légèrement le taux d'erreur, `maxConf` diminue un peu plus le taux d'erreur, et les meilleurs résultats sont donnés par `maxConfSupp`.

Le bénéfice de chaque stratégie est lié à la taille de la base de cas : plus la base de cas contient de cas pour l'acquisition des compétences, meilleurs sont les résultats. Comparé à la stratégie de référence, le bénéfice pour la meilleure stratégie pour la sélection `maxConfSupp` est remarquable. Même si le taux de réponse est déjà très bon, avec la stratégie de référence et d'autant plus avec un taux de réponse de 100%, `maxConfSupp` l'améliore encore en se rapprochant significativement de 100% de réponses correctes. Pour `DNF`, selon la taille de la base de cas (64, 96 et 128), le taux d'erreur `%err` passe de 10,1 à 6,9 (diminution de 32%), de 8,4 à 3,3 (diminution de 61%) et de 7,7 à 1,7 (diminution de 78%). Pour `Pol`, les résultats sont encore plus impressionnant, selon la taille de la base de cas (64, 96 et 128), le taux d'erreur `%err` passe de 13,7 à 6,3 (diminution de 54%), de 10,5 à 1,6 (diminution de 859%) et de 8,8 à 0,5 (diminution de 94%).

Ainsi, les premiers résultats expérimentaux montrent que, à partir d'une certaine taille de la base de cas, la stratégie `maxConfSupp` surpasse toutes les autres méthodes et diminue nettement le taux d'erreur, tout en utilisant un nombre de triplets plus restreint.

5 Discussion et travaux proches

Dans cette section, l'approche présentée dans ce papier est comparée à la littérature sur le RàPC selon deux angles : la notion de compétence et l'apprentissage de connaissances d'adaptation.

Les compétences en RàPC. Dans [13], trois types de processus de RàPC sont distingués, en particulier l'extrapolation, qui remémore et réutilise les cas par triplets, et l'approximation, qui remémore et réutilise les cas par singletons. Par rapport aux recherches précédentes sur la compétence qui sont liées à l'approximation, les travaux présentés dans cet article considère une notion de compétence liée à l'extrapolation.

La notion de compétence en RàPC est utilisée en général pour la maintenance de la base de cas, soit pour supprimer les cas les moins compétents [17] (dans l'optique de minimiser la perte de compétence globale), soit ajouter des cas compétents [19] (dans l'optique de maximiser le gain de compétence globale). Dans ces précédentes études, la compétence est liée au cas sources, individuellement, par rapport aux autres cas de la base de cas. En particulier, dans l'article fondateur [17], la compétence des cas est évaluée en mettant les cas sources dans des catégories (des cas centraux — *pivotal cases* —, qui sont les plus compétents, aux cas auxiliaires), ces catégories étant définies avec l'aide d'une notion binaire de l'adaptabilité entre cas et problème. Ainsi, la notion de compétence est liée au processus d'approximation, considérant les cas sources individuellement.

En revanche, cet article aborde la compétence liée au processus d'extrapolation : les cas sont remémorés par triplets. La compétence d'un triplet $(a, b, c) \in \mathtt{BC}^3$ est réduite à la compétence du couple $(a, b) \in \mathtt{BC}^2$, qui est liée à l'ensemble des autres couples $(c, d) \in \mathtt{BC}^2$. Un point commun à ces deux notions de compétence est que la compétence d'un objet (l'objet étant un cas pour l'approximation et un couple de cas pour l'extrapolation) n'est pas une propriété intrinsèque à l'objet, mais est liée aux autres objets (de $\mathtt{BC}$ ou $\mathtt{BC}^2$).

Une différence mineure entre les études précédentes sur la compétence et celle-ci est liée à l'utilisation de la com-

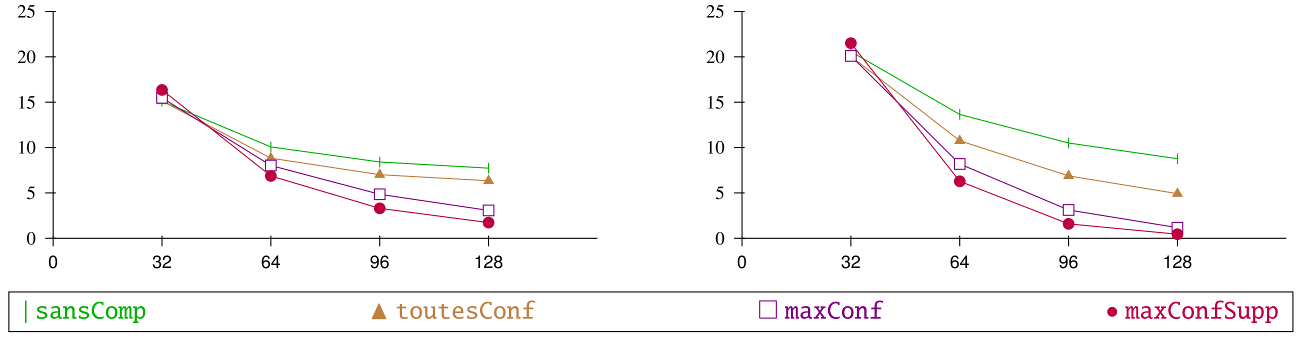

Figure 1 – Taux d'erreur en fonction de la taille de BC, pour chaque générateur (à gauche : DNF, à droite : Pol).

		$\|BC\| = 32$		$\|BC\| = 64$		$\|BC\| = 96$		$\|BC\| = 128$	
		%err	%rep	%err	%rep	%err	%rep	%err	%rep
DNF	sansComp	**15,1**	97,5	10,1	100,0	8,4	100,0	7,7	100,0
	toutesConf	15,1		8,8		7,0		6,3	
	maxConf	15,5		8,0		4,9		3,1	
	maxConfSupp	16,4		**6,9**		**3,3**		**1,7**	
POL	sansComp	20,6	95,8	13,7	100,0	10,5	100,0	8,8	100,0
	toutesConf	**20,1**		10,8		6,9		4,9	
	maxConf	20,1		8,2		3,1		1,2	
	maxConfSupp	21,5		**6,3**		**1,6**		**0,5**	

Table 2 – %err et %rep pour les différentes stratégies de sélection et vote pour les différents générateurs.

pétence : la maintenance de la base de case pour les précédentes et la résolution de problème pour celle-ci.

Liens avec l'apprentissage de connaissances d'adaptation. Le travail présenté dans cet article a des liens forts avec la majorité des travaux sur l'apprentissage de connaissances d'adaptation (ACA). L'adaptation dont il est question ici est celle qui suit la remémoration d'un seul cas (c'est une partie d'un processus d'approximation intégrant une adaptation). Elle s'appuie en particulier sur les connaissances d'adaptation AK qui peuvent se définir informellement par :

$$\mathtt{AK} = \text{« Comment varie la solution quand varie le problème. »}$$

La démarche très généralement suivie en ACA, initiée par le travail de Kathleen Hanney et Mark T. Keane [9], s'appuie sur la base de cas pour extraire des connaissances d'adaptation selon le principe suivant. Elle considère un ensemble EA de couples de cas sources (a, b), avec $a \neq b$, soit en prenant tous les couples distincts de BC, soit en se restreignant sur la base d'une similarité minimale entre a et b. Puis, EA est utilisé comme ensemble d'apprentissage d'un processus d'apprentissage supervisé : pour tout couple (a, b), l'entrée est le couple $(\mathtt{x}^a, \mathtt{x}^b)$ et la sortie est le couple $(\mathtt{y}^a, \mathtt{y}^b)$. Le processus d'apprentissage supervisé donne alors un modèle de cette connaissance AK, utilisée par l'adaptation.

Plusieurs travaux s'inscrivent dans ce schéma général. Dans [10], AK consiste en une représentation de « cas d'adaptation ». Dans [5], différentes techniques sont utilisées, en particulier l'induction par arbre de décisions et des techniques d'apprentissage ensembliste/ Dans [6], l'extraction de motifs fermés fréquents est utilisée. L'interprétation par des experts du domaine d'application permet de produire des règles d'adaptation à ajouter à AK. Dans [8], des techniques similaires à celles de [6] sont utilisées (l'analyse formelle de concepts et l'extraction de motifs fermés fréquents étant des techniques de fouille de données très proches), mais, dans ce travail, des cas négatifs (i.e., des couples $(\mathtt{x}, \mathtt{y}) \in \mathcal{P} \times \mathcal{S}$ pour lesquels $\mathtt{y}$ n'est pas une solution correcte de $\mathtt{x}$) sont utilisés, ce qui améliore nettement les résultats de l'apprentissage.

On peut considérer que le travail présenté dans cet article relève également de l'ACA. En fait, nous avons déjà utilisé la notion de règle d'adaptation pour considérer un couple (a, b) de cas ; précisons cette idée. Soit la relation $\sim$ définie, pour (a, b) et (a', b'), deux couples de cas, par :

$$(a, b) \sim (a', b') \quad \text{si} \quad \mathtt{x}^a : \mathtt{x}^b :: \mathtt{x}^{a'} : \mathtt{x}^{b'} \text{ et } \mathtt{y}^a : \mathtt{y}^b :: \mathtt{y}^{a'} : \mathtt{y}^{b'}$$

Pour les proportions analogiques sur les représentations nominales définies en section 2.2, $\sim$ est une relation d'équivalence [4]. Donc, résoudre un problème $\mathtt{x}^{\mathtt{cible}}$ par extrapolation à partir d'un triplet $(a, b, c) \in \mathtt{BC}^3$ ou d'un triplet $(a', b', c) \in \mathtt{BC}^3$ (avec le même c) tels que $(a, b) \sim (a', b')$, donnera le même résultat : le choix du représentant de la classe d'équivalence de (a, b) pour $\sim$ est indifférent pour l'extrapolation. On peut utiliser une telle classe d'équivalence C comme une règle d'adaptation (où c est le cas remémoré et $\mathtt{x}^{\mathtt{cible}}$ est le problème à résoudre) :

avec (a, b) choisi arbitrairement dans C

si $\mathtt{x}^a : \mathtt{x}^b :: \mathtt{x}^c : \mathtt{x}^{\mathtt{cible}}$ et $\mathtt{y}^a : \mathtt{y}^b :: \mathtt{y}^c : y$ a une solution

alors cette solution est une solution plausible de $\mathtt{x}^{\mathtt{cible}}$

L'ensemble des classes d'équivalences de la restriction de $\sim$ à $\mathtt{BC}^2$ donne donc un ensemble de règles d'adaptation candidates, mais toutes les règles ne se valent pas : certaines sont plus plausibles que d'autres. Un critère doit donc être défini pour effectuer une préférence entre ces règles ou, si on décide de les appliquer toutes, de le faire en pondérant dans un vote certaines règles.

Une façon simple de faire cela (utilisée, par exemple, dans [6]) consiste à utiliser le cardinal de C. On rejoint là l'idée de compétence d'un couple de cas : si $(a, b) \in C$, alors, $|C| = \mathtt{supp}(a, b) \times \mathtt{conf}(a, b)$. La limite de cette approche est qu'elle ne compte que les exemples (à l'appui de la règle), pas les contre-exemples (qui pénalisent la règle). En revanche, l'approche considérée dans ce papier tient compte des contre-exemples : si on sait, par exemple, que $\mathtt{conf}(a, b) = 1/3$, alors on sait que pour chaque exemple, il y a deux contre-exemples, donc même si $\mathtt{supp}(a, b)$ est important, la règle associée à (a, b) sera douteuse.

Une autre différence avec le travail de [6] est que dans [6], on considérait des motifs de variations qui peuvent être approximatifs. Par exemple, si (a, b) et (a', b') sont deux couples de cas sources tels que, pour *la plupart* des attributs i, $\mathtt{x}_i^a : \mathtt{x}_i^b :: \mathtt{x}_i^{a'} : \mathtt{x}_i^{b'}$, on considérera la règle construite sur ces attributs communs, en négligeant le reste. Dans un cadre où les analogies sont rares (par exemple, quand on a des attributs à valeurs réelles), remplacer la proportion analogique exacte par une proportion analogique graduelle [7] dans l'approche décrite dans cet article pourrait se justifier, ce qui constitue une perspective de recherche potentielle de ce travail. En effet, la démarche présentée dans cet article pourrait s'appliquer avec une analogie graduelle, en particulier avec des attributs numériques.

Cette discussion montre que certaines idées d'ACA à partir de la base de cas se trouvent naturellement reformulées dans le cadre des proportions analogiques : ce lien établi entre les deux champs est donc potentiellement fécond.

4. La réflexivité et la symétrie découlent des postulats éponymes, en revanche, la transitivité n'est pas réalisée par toutes les proportions analogiques [12].

6 Conclusion

Le RàPC classique s'appuie sur les similarités individuelles du problème à résoudre avec chacun des problèmes résolus déjà connus. Nous avons montré qu'il pouvait également être intéressant d'envisager des triplets de cas (a, b, c) pour rendre égaux le changement de a à b avec le changement de c au problème à résoudre associée à sa solution potentielle. C'est le fondement de l'extrapolation analogique fondée sur les proportions analogiques. Cependant, il a été observé que certains triplets peuvent conduire à de mauvaises inférences.

Dans cet article, nous avons proposé de discriminer les triplets par une évaluation de la « compétence » de couples participant aux triplets. En effet, une proportion analogique « a est à b ce que c est à d » peut être vue comme l'établissement d'un parallèle entre ces deux couples. Les différences entre les attributs d'un couple de problèmes sont naturellement liées aux différences entre leurs solutions, mais cette relation dépend du contexte exprimé par les valeurs des attributs qui ne changent pas. Nous avons montré qu'il était possible, au moins dans une certaine mesure, d'évaluer la compétence de couples pour sélectionner les « bons » triplets et améliorer les résultats de l'inférence analogique. Ceci contribue à confirmer l'intérêt de l'extrapolation analogique pour le RàPC.

Une première perspective de ce travail a été évoquée à la fin de la section 3. Elle consiste, lors du choix d'un triplet (a, b, c), à considérer non seulement la compétence de (a, b) mais également celle de (a, c).

Une autre perspective de ce travail consiste à s'intéresser à l'utilisation d'idées dans le champ du RàPC consacré à l'acquisition de connaissances d'adaptation (cf. section 5). En particulier, on pourrait envisager d'améliorer l'approche décrite dans ce papier en s'appuyant sur l'utilisation d'une base de cas négatifs, comme c'est le cas dans [8].

Enfin, on peut également s'intéresser à la compétence pour associer à la solution proposée par le système de RàPC une indication sur sa plausibilité, qui sera d'autant meilleure que la compétence sera grande. L'intérêt de cela serait de voir comment combiner l'extrapolation utilisant la compétence avec d'autres approches du RàPC qui donne également une indication de plausibilité.

Références

[1] Aamodt, A. et E. Plaza: *Case-based Reasoning : Foundational Issues, Methodological Variations, and System Approaches*. AI Communications, 7(1) :39–59, 1994.

[2] Bounhas, M., H. Prade et G. Richard: *Analogical classification : A rule-based view*. Dans *Proc. of the International Conference on Information Processing*

and Management of Uncertainty in Knowledge-Based Systems, pages 485–495. Springer, 2014.

[3] Correa, W. F., H. Prade et G. Richard: *Trying to understand how analogical classifiers work*. Dans *Proc. of the International Conference on Scalable Uncertainty Management*, pages 582–589. Springer, 2012.

[4] Couceiro, M., N. Hug, H. Prade et G. Richard: *Analogy-preserving functions : a way to extend Boolean samples*. Dans *Proc. of the 26th Int. Joint Conf. on Artificial Intelligence (IJCAI'17)*, pages 1575–1581. Morgan Kaufmann, Inc., 2017.

[5] Craw, S., N. Wiratunga et R. C. Rowe: *Learning adaptation knowledge to improve case-based reasoning*. Artificial Intelligence, 170(16-17) :1175–1192, 2006.

[6] d'Aquin, M., F. Badra, S. Lafrogne, J. Lieber, A. Napoli et L. Szathmary: *Case base mining for adaptation knowledge acquisition*. Dans Veloso, M. M. (rédacteur) : *Proc. of the 20th Int. Joint Conf. on Artificial Intelligence (IJCAI'07)*, pages 750–755. Morgan Kaufmann, Inc., 2007.

[7] Dubois, D., H. Prade et G. Richard: *Multiple-valued extensions of analogical proportions*. Fuzzy Sets and Systems, 292 :193–202, 2016. `https://doi.org/10.1016/j.fss.2015.03.019`.

[8] Gillard, T., J. Lieber et E. Nauer: *Improving Adaptation Knowledge Discovery by Exploiting Negative Cases : First Experiment in a Boolean Setting*. Dans *Proc. of ICCBR 2018 - 26th International Conference on Case-Based Reasoning*, Stockholm, Sweden, juillet 2018. `https://hal.inria.fr/hal-01905077`.

[9] Hanney, K. et M. T. Keane: *Learning adaptation rules from a case-base*. Dans Smith, I. et B. Faltings (rédacteurs) : *Advances in Case-Based Reasoning – Proc. of the Third Eur. Workshop, EWCBR'96*, LNAI 1168, pages 179–192. Springer Verlag, Berlin, 1996.

[10] Jarmulak, J., S. Craw et R. Rowe: *Using case-base data to learn adaptation knowledge for design*. Dans *Proc. of the 17th Int. Joint Conf. on Artificial Intelligence (IJCAI'01)*, pages 1011–1016. Morgan Kaufmann, Inc., 2001.

[11] Kolodner, J.: *Case-Based Reasoning*. Morgan Kaufmann, Inc., 1993.

[12] Lepage, Y.: *Proportional analogy in written language data*. Dans Gala, N., R. Rapp et G. Bel-Enguix (rédacteurs) : *Language, Production, Cognition and the Lexicon*, Text, Speech and Language Technology 48, pages 151–173. Springer International Publishing Switzerland, 2014. `https://link.springer.com/chapter/10.1007/978-3-319-08043-7_10`.

[13] Lieber, J., E. Nauer, H. Prade et G. Richard: *Making the Best of Cases by Approximation, Interpolation and Extrapolation*. Dans *Proc. of ICCBR 2018 - 26th International Conference on Case-Based Reasoning*, tome 11156, pages 580–596, Stockholm, Sweden, juillet 2018. Springer. `https://hal.inria.fr/hal-01905058`.

[14] Prade, H. et G. Richard: *A discussion of analogical-proportion based inference*. Dans Sánchez-Ruiz, A. A. et A. Kofod-Petersen (rédacteurs) : *Proc. ICCBR'17 Workshops (CAW, CBRDL, PO-CBR), Doctoral Consortium, and Competitions co-located with the 25th Int. Conf. on Case-Based Reasoning (ICCBR'17), Trondheim, June 26-28*, tome 2028 de *CEUR Workshop Proceedings*, pages 73–82, 2017.

[15] Prade, H. et G. Richard: *Analogical proportions : From equality to inequality*. Int. J. Approx. Reasoning, 101 :234–254, 2018.

[16] Richter, M. M. et R. O. Weber: *Case-based reasoning, a textbook*. Springer, 2013.

[17] Smyth, B. et M. T. Keane: *Remembering to forget*. Dans *Proc. of the 14th Int. Joint Conf. on Artificial Intelligence (IJCAI'95)*, Montréal, 1995.

[18] Stroppa, N. et F. Yvon: *Analogical learning and formal proportions : Definitions and methodological issues*. Technical Report D004, ENST-Paris, 2005.

[19] Zhu, J. et Q. Yang: *Remembering to add : competence-preserving case-addition policies for case base maintenance*. Dans *Proc. of the 16th Int. Joint Conf. on Artificial Intelligence (IJCAI'99)*, pages 234–241, 1999.